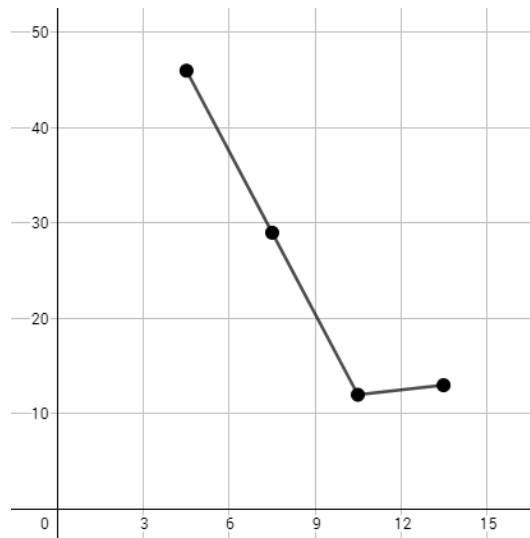
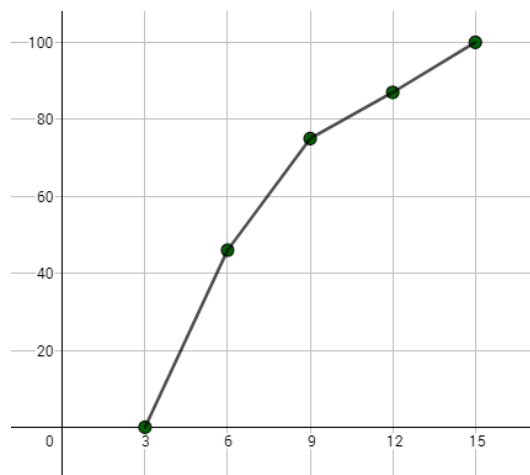


Exercise 5A:

1.



a.



b.

c. 4m 38s

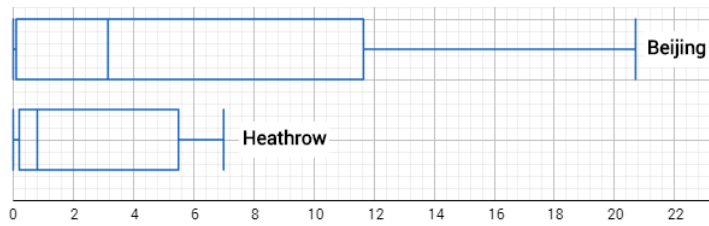
Exercise 5B:

1.

- Ease of collection
- Not everyone asked may have had a dog that had died / some people may have refused to respond
- Median = 5, IQR = 8
- Cats live longer on average but have a smaller spread of lifespans
- The sample sizes are dissimilar – would need to collect more data about dogs

2.

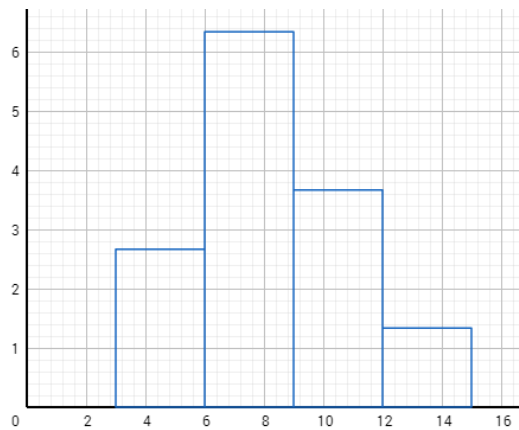
- a. Beijing
- b. "Trace" – less than 0.05mm of rain



- c.
- d. Beijing receives more rain, on average, than Heathrow, and has a greater spread of daily rainfall totals.

Exercise 5C:

1.



- a.
 - b. 25 children
- 2.
- a. Continuous data that has been grouped
 - b. Area of bar is proportional to frequency
 - c. Width = $1\frac{2}{3}$ cm, Height = 4.74cm
 - e. £51,984

3.

- a. There are 234 weighted squares in total and there are 39 squares that represent a score between 70–80, so number of students are

$$\frac{210}{234} \times 39 = 35$$

- b. We have:

$$\begin{aligned}\bar{x} &\approx \frac{210}{234} \frac{[(10 \times 12) + (40 \times 40) + (65 \times 38) + \dots + (140 \times 2)]}{210} \\ &= 76.88 = 76.9\end{aligned}$$

- c. We first look for the class interval that contains the 105th student (by calculating cumulative frequencies).

We find that for 0–70, the cumulative frequency is 79.079..., and for 0–90, it is 147.615... .

Hence, the median lies in the class 70–90.

Now, we can use interpolation to find the median.

$$\begin{aligned}\frac{147.615\dots - 79.079\dots}{90 - 70} &= \frac{147.615\dots - 105}{90 - Q_2} \\ \Rightarrow Q_2 &= 77.6\end{aligned}$$

Exercise 5D:

1.

- a. 56
b. 9.58
c. 10.4
d. $1.5 \times IQR = 1.5 \times 10.4 = 15.6$

$15.6 + 9.58 = 25.18$. All the data is less than 25.18, so there are no ('upper') outliers.

$9.58 - 15.6 = -6.02$. All the data is greater than -6.02 (in particular, the values are non-negative), so there are no ('lower') outliers.

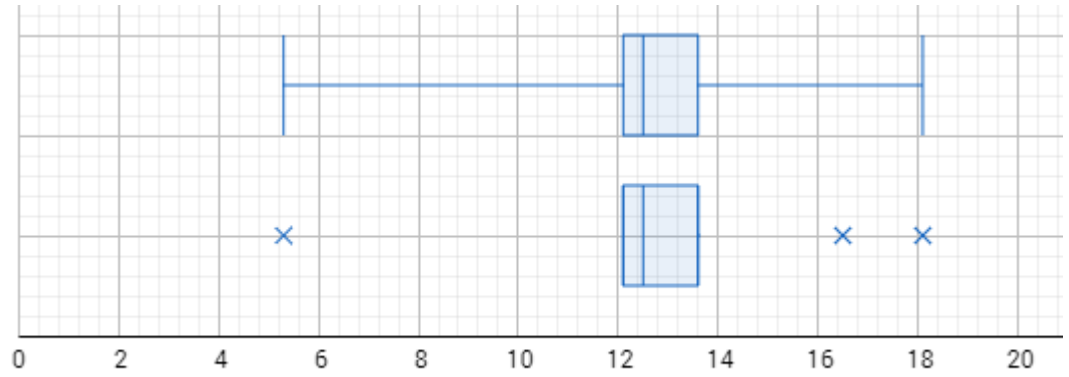
Hence, the data has no outliers.

2.

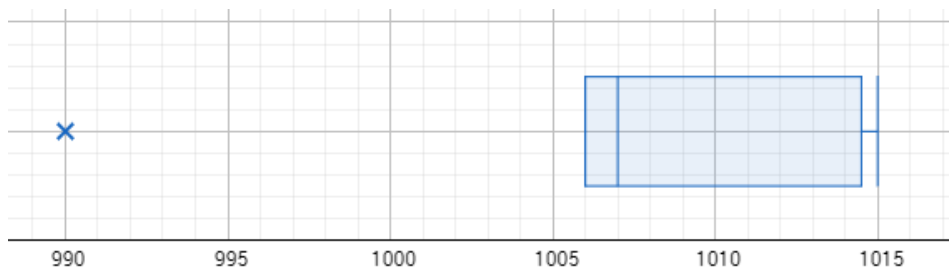
a. Median = 12.5, IQR = 1.5

b. Mean = 12.77, SD = 3.02

c.



3.



Exercise 5E:

2.

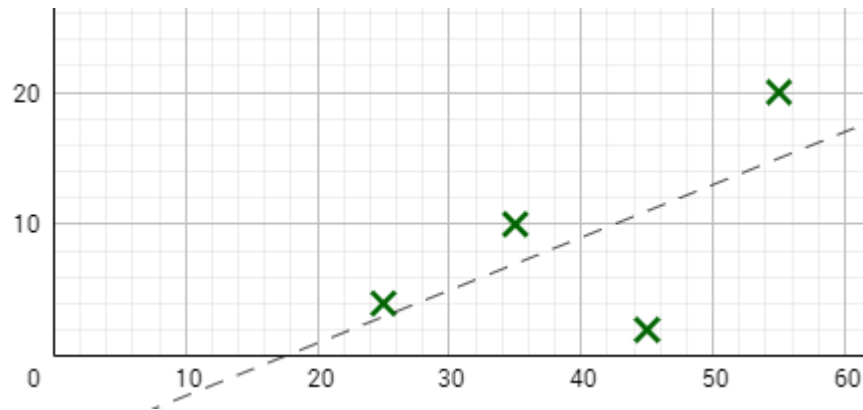
a. Strong positive correlation

b. Gradient ≈ 17 , Intercept ≈ -112

c. £324.30

d. Outside of given range (extrapolation) so may not be reliable

3.



a.

(Moderate) positive correlation

b.

i. 45 lux

ii. 15 bubbles/min

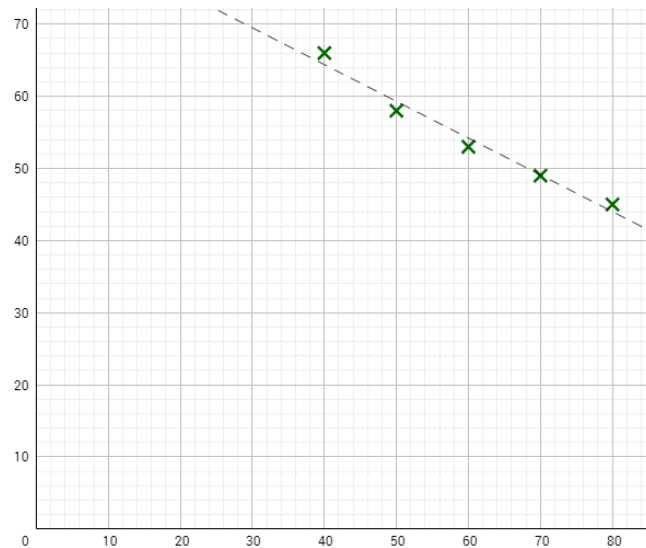
c. Reduces the strength of correlation

d. Assuming correlation implies causation

e. Scientific reasoning/an explanation

4.

a. Speed is the explanatory variable



b.

c. Very strong negative correlation: the faster the speed of David's car, the lower the fuel consumption

d. 56.75mpg

- e. Extrapolation – falls outside of the range of given data. It is also unlikely that David's car would be able to reach 200mph, nor would it be legal to test this under most circumstances.